

## Los Repositorios Federados de Institución, una propuesta *de articulación de la vía verde y la vía dorada del Acceso Abierto en ALyC*

En el Acceso Abierto (*Open Access*, OA) se han trazado dos vías de implementación; por un lado, la vía dorada referente a la publicación en revistas arbitradas y la vía verde correspondiente al autoarchivo en repositorios institucionales.

La intersección de conjuntos que suponen las dos rutas habilita el archivo de artículos publicados en revistas arbitradas de acceso abierto en repositorios institucionales. Sin embargo, los modelos de *copyright* con los que operan las revistas científicas o la falta de claridad en la definición de los límites y posibilidades del autoarchivo y la publicación en revistas, ha propiciado una marcada ausencia de artículos *postprint* en los repositorios de las instituciones.

Por definición un Repositorio Institucional (RI) es un sistema de información que reúne, preserva, divulga y da acceso a la producción intelectual y académica de una institución.

Si bien el desarrollo e implementación de repositorios por parte de las instituciones en América Latina y el Caribe mantiene una tendencia a la alza, se observa (en registros como ROAR) que los contenidos depositados en ellos son principalmente tesis, objetos de enseñanza-aprendizaje, recursos educativos, imágenes, publicaciones editadas por las instituciones, videos; inclusive, fondo antiguo. Sin

embargo, es clara la casi nula presencia de *postprints* en los RI's, es decir, las instituciones no tienen en sus repositorios la producción de sus investigadores publicada en revistas arbitradas.

En el caso de México, por ejemplo, más del 70% de los contenidos de los repositorios de las instituciones registradas en la Red Mexicana de Repositorios Institucionales corresponde a tesis, imágenes y videos; y aunque dicha red registra 29% de artículos, estos son obtenidos de repositorios de revistas, artículos producidos por diversas instituciones y publicados en esas revistas. Es decir, los repositorios de las instituciones mexicanas hasta el momento no mantienen un archivo representativo de artículos *postprint* producidos por ellas.

Es así, como se visualiza una práctica casi nula de autoarchivo de artículos revisados por pares en los repositorios de las instituciones de la región.

Por otro lado, hablando de la vía dorada del OA, los autores publican en una revista de acceso abierto que provee disponibilidad de sus artículos a través de su sitio web o de portales de revistas. Estos portales de revistas, como Scielo o Redalyc, funcionan a su vez, como repositorios para la vía dorada y concentran los contenidos de las publicaciones científicas de acceso abierto.

Repositorios cross o multi-institucionales y repositorios temáticos también concentran trabajos resultado de investigaciones de diversas instituciones, tal es el caso de ArXiv, RePEc, entre otros.

Si los datos y contenidos de los artículos de estos repositorios son correctamente tratados y sometidos a procesos de normalización es posible agrupar y organizar los artículos de acuerdo a la institución de afiliación de sus autores.

De ese modo, estas fuentes de información son capaces de proveer un "Repositorio Federado" para cada institución con su producción científica. Un

“Repositorio Federado de Institución” o RFI adquiere relevancia al implementar protocolos de interoperabilidad de tal forma que los metadatos que contienen puedan ser integrados en los repositorios institucionales. Así un RI reúne la producción de su institución que ha sido publicada en revistas arbitradas y la cual no ha podido ser depositada por parte de los autores, por las circunstancias expuestas.

Los “Repositorios Federados” para una institución fungen como el complemento de un Repositorio Institucional para la producción científica publicada en revistas arbitradas, estos deben cumplir con un protocolo de interoperabilidad como OAI-PMH.

La propuesta de los RFI's se materializa en Redalyc.org con la generación de micrositios de más de 9,000 instituciones de América Latina y el Caribe con artículos publicados por sus investigadores y que a la vez ofrecen servicios de búsqueda, navegación e interconexión a través de OAI-PMH, lo que además les permite interoperar con los repositorios implementados en las instituciones.

A través de los RFI's, se busca regresar ese valor depositado en las publicaciones de la vía dorada del OA a la institución que las produjo.

## 1. La complementariedad en las vías del Acceso Abierto

Varios canales diferentes para comunicar los resultados de investigación seguirán siendo importantes en los próximos años. Sin embargo, el acceso abierto es el futuro de la publicación académica. (Finch, 2012)

Las vías dorada y verde del *Acceso Abierto (AA)* son dos modelos diferentes que se traslapan; es decir, la publicación en revistas de acceso abierto y el depósito en

repositorios, respectivamente, duplican la difusión en línea de los artículos *postprint* y se puede ilustrar con la teoría de conjuntos como en la figura 1.

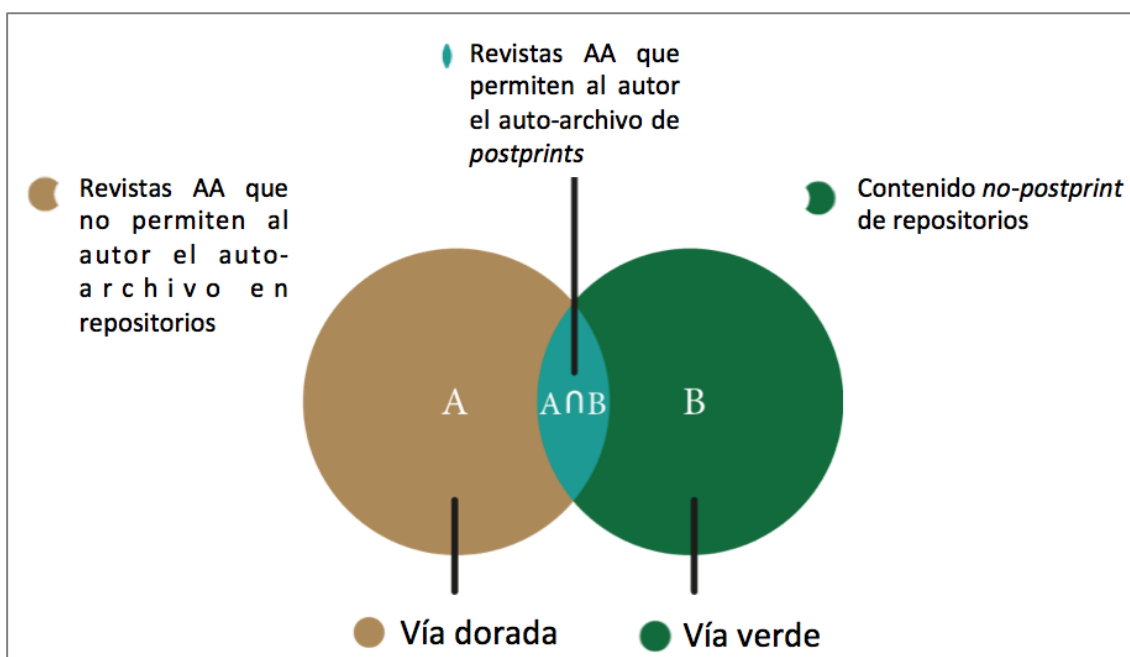


Figura 1. Complementariedad de la vía dorada y la vía verde del AA

La ruta verde usa los repositorios en línea para almacenar duplicados digitales de resultados de investigación publicados y para ponerlos disponibles abierta y libremente a través del internet. Hay que notar que este no es un reemplazo de la publicación tradicional, o de la publicación de acceso abierto, pero existe más allá de la publicación como una forma suplementaria de diseminación. (Hubbard, 2012)

En un proceso de publicación tradicional los autores ceden el *copyright* a la revista. Esto significa que mientras las revistas no retornen los derechos a los autores, estos son incapaces de disponibilizar su trabajo en internet. Algunas revistas permiten al autor únicamente el archivo en línea del *preprint*, es decir, del artículo borrador sin correcciones. Esta situación impide que el autor deposite el *postprint* –versión final del autor o versión publicada por la revista- en los repositorios de la institución auspiciadora de la investigación.

Los llamados editores verdes -que permiten el autoarchivo de *preprints* y *postprints*- o los azules –que autorizan el autoarchivo de *postprints*- son los que favorecen la población de contenidos en los repositorios; sin embargo, la proliferación de revistas que operan con este modelo es baja. Consecuencia de lo anterior es el escaso registro de *postprints* en los repositorios de las instituciones.

Si a ello se le suma que los repositorios, según (Abadal, 2013) y el informe STM (Ware, 2012), crecen en número pero no en contenidos y que se tiene una visión escéptica sobre su evolución; se vislumbra un panorama no alentador en la consolidación de la vía verde para la difusión de los resultados publicados en revistas científicas.

## 2. Los Repositorios Institucionales en ALyC

Se entiende como Repositorio Institucional (RI) un sistema de información que reúne, preserva, divulga y da acceso a la producción intelectual y académica de una institución. Los contenidos de un RI varían entre tesis, objetos de aprendizaje, imágenes, videos, libros, capítulos de libro, revistas, artículos, etc.

Para el caso, de los artículos, en la vía verde del acceso abierto los autores publican en una revista y posteriormente depositan una versión del artículo para su libre acceso en un repositorio institucional, lo que se espera que sea depositado es el *postprint*. Sin embargo, como se mencionó en el apartado anterior, esta condición solo es posible si las revistas permiten que los autores ejerzan el autoarchivo de sus artículos en los repositorios de sus instituciones.

Teniendo en consideración que la base de datos de Redalyc, que hoy en día cuenta con más de 800 revistas, contiene una muestra muy representativa de las revistas de Acceso Abierto en ALyC y que menos de un 5% de las revistas que indiza

operan bajo un modelo de “revista verde” claramente definido, es posible afirmar que el depósito de *postprints* en repositorios institucionales no es una práctica común en la región.

Ahora bien, se ha mostrado un comportamiento generalizado en América Latina y el Caribe acerca de la población de repositorios institucionales principalmente con contenidos tales como tesis, objetos de enseñanza-aprendizaje, recursos educativos, imágenes, publicaciones editadas por las instituciones, videos; inclusive, fondo antiguo. Sin embargo se destaca la casi nula presencia de *postprints* en los RIs, es decir, las instituciones no tienen en sus repositorios la producción de sus investigadores publicada en revistas arbitradas.

La figura 3 muestra los repositorios existentes en ALyC, cabe destacar que fuera de Brasil el establecimiento de repositorios en cada país no supera los 35, según el Registro de Repositorios de Acceso Abierto (ROAR).

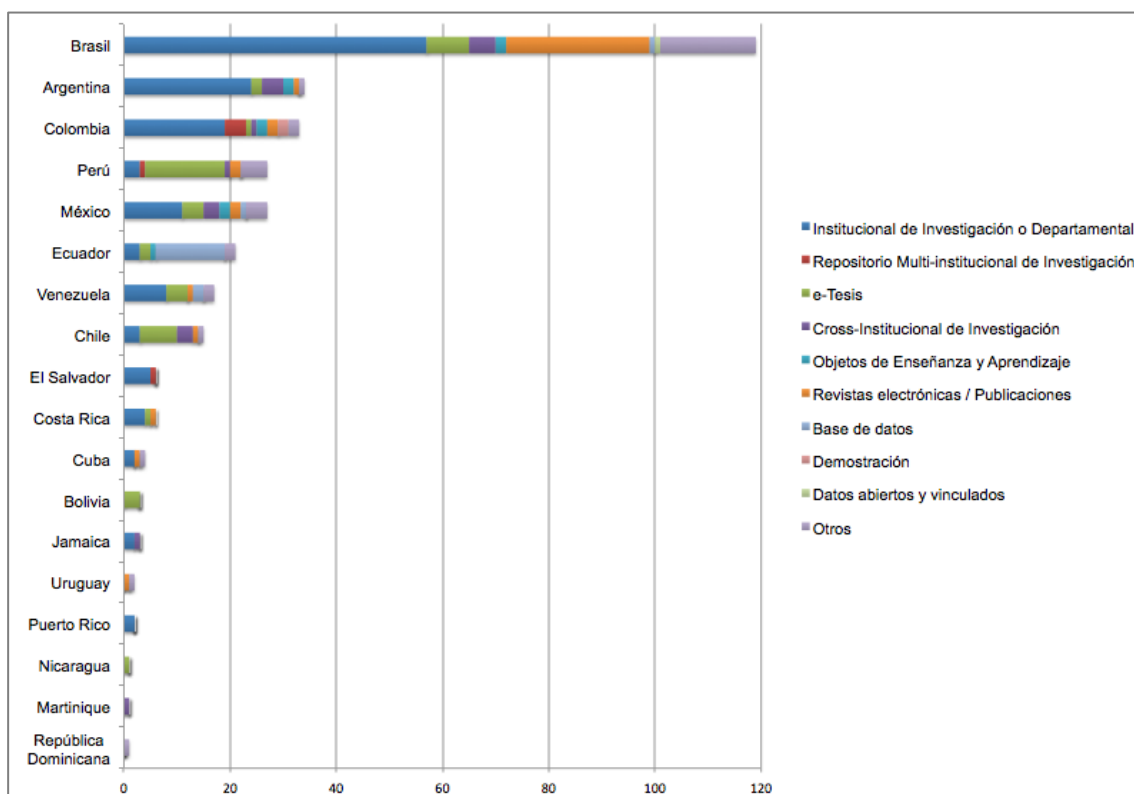


Figura 2. Repositorios por país en América Latina en el Registro de Repositorios de Acceso Abierto.

(ROAR)

Partiendo de la clasificación por tipo de repositorio, se observa que los repositorios institucionales de investigación o departamentales son los que pueden contener artículos *postprint* generados por autores adscritos a la institución que hospeda el repositorio, ya que el resto por definición almacenan otros tipos de contenidos como es el caso de tesis, bases de datos y objetos de enseñanza aprendizaje.

Los llamados cross-institucionales y los multi-institucionales, es posible que contengan artículos *postprint*, sin embargo, su procedencia es de diversas instituciones por lo que no se puede asociar esos contenidos como producción de la institución que mantiene el repositorio.

Para el caso de los repositorios de revistas electrónicas o publicaciones sucede una circunstancia similar. Generalmente, son repositorios de revistas que edita la institución y que por lo tanto contemplan trabajos que son resultados de investigaciones de otras instituciones.

En la figura 4, se ilustra en azul la proporción de repositorios institucionales de investigación o departamentales, espacios posibles de depósito de artículos *postprint*. Como se puede observar el espectro está dominado por repositorios de tesis y de revistas electrónicas.

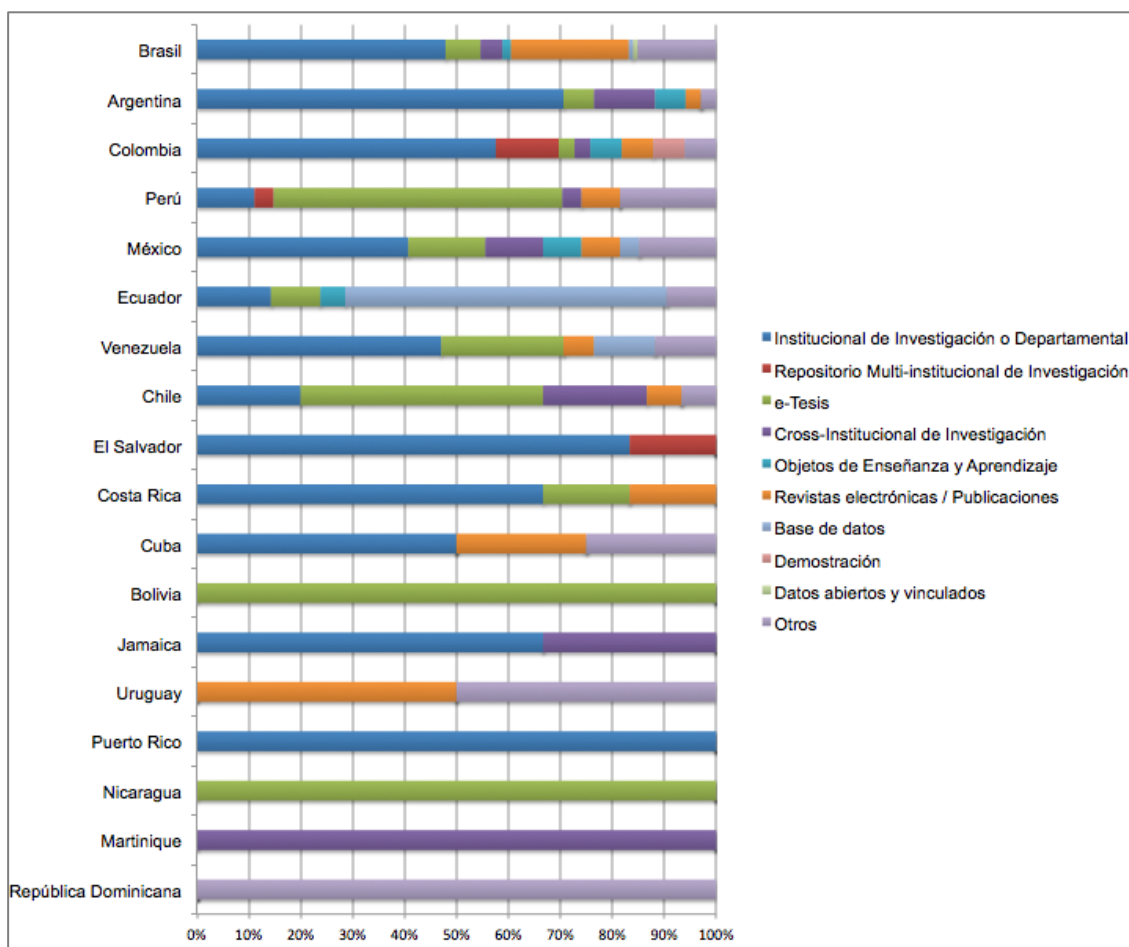


Figura 3. Proporción por tipo de repositorio en los países de América Latina y el Caribe en el Registro de Repositorios de Acceso Abierto. (ROAR)

En México por ejemplo, el diagnóstico nacional de repositorios institucionales realizado en 2012 para las Instituciones de Educación Superior (IES) mexicanas (Rodríguez, 2013) mostró que de 55 IES participantes tan solo el 38% cuentan con uno o más repositorios. Como se observa en la figura 2, los contenidos más comunes son: tesis y disertaciones con un 21%, después libros y capítulos de libro con un 18% y solo 15% de los repositorios cuentan con artículos *preprint* o *postprint*.



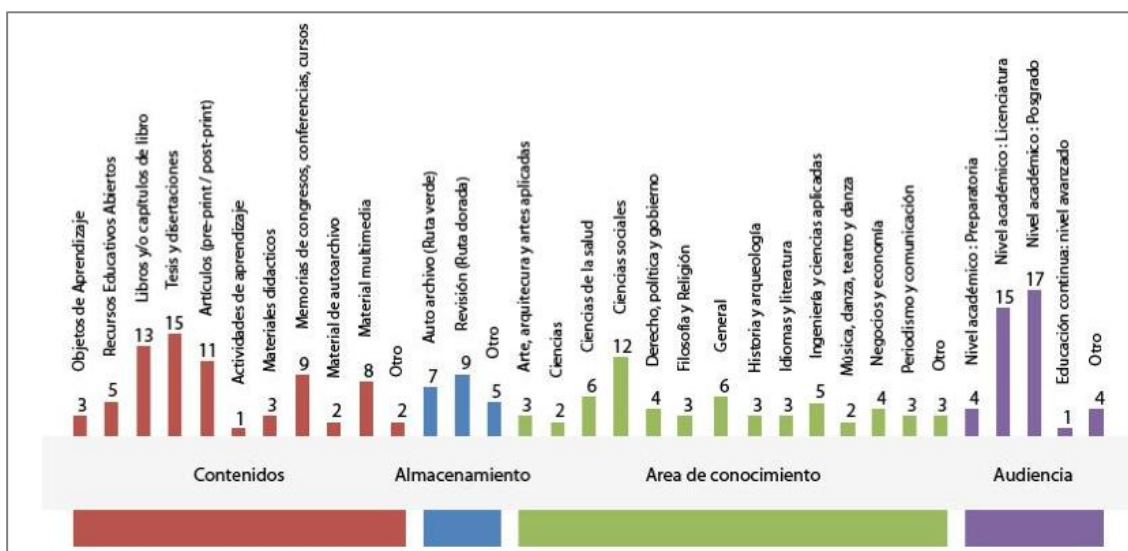


Figura 4. Contenidos de los repositorios, vía de almacenamiento, áreas de conocimiento y audiencia de los repositorios institucionales mexicanos. Fuente **(Rodríguez, 2013)**

Por su parte, la recién formada Red Mexicana de Repositorios Institucionales – REMERI- (REMERI, 2012) actualmente integra 18 repositorios de 12 instituciones mexicanas reuniendo un total de 106,169 trabajos entre artículos, tesis, libros, revistas, videos, imágenes; distribuidos como se muestra en la tabla 1.

Tipo de Material	Cantidad de Registros
Tesis de Licenciatura	16.6%
Tesis de Maestría	4.3%
Tesis de Doctorado	0.1%
Imágenes	47.0%
Artículos	29.2%
Videos	2.9%

Tabla 1. Proporción por tipo de contenido de REMERI. **(REMERI, 2012)**

Se puede observar que son imágenes lo que principalmente aportan los repositorios mexicanos a la red, por su parte 29.2% de los depósitos corresponde a artículos; sin embargo, al explorar el metabuscador de la red se puede corroborar que dichos artículos son provistos por repositorios de revistas editadas por instituciones

(Portal de Revistas Científicas de la UAEH y Portal de Revistas Científicas y Arbitradas de la UNAM).

Lo que quiere decir que si bien son artículos científicos, estos no corresponden a resultados de investigaciones producidas por la UNAM o UAEH exclusivamente, ya que las revistas publican manuscritos de diversas instituciones.

### 3. Los Repositorios Federados de Institución, la propuesta para retornar la producción dorada a una institución.

Tomando lo anterior en consideración, entonces ¿cómo puede una institución disponibilizar en línea su producción publicada en revistas de AA?.

No hay que olvidar que tras la publicación en una revista de acceso abierto aparece un actor más en los procesos de diseminación científica y son los repositorios y/o portales de revistas, como Scielo o Redalyc, los cuales fungen a su vez, como repositorios para la vía dorada y concentran los contenidos de las publicaciones científicas de acceso abierto.

Si los datos y contenidos de los artículos publicados en revistas OA y disponibles en línea a través de dichos portales son correctamente tratados y sometidos a procesos de normalización es posible clasificar los artículos de acuerdo a la institución de afiliación de sus autores.

De este modo los portales de revistas se convierten en plataformas esenciales capaces de proveer un “repositorio” para cada institución con su producción científica publicada en revistas de acceso abierto indizadas en estas bases. Estos almacenes

pueden generar “repositorios federados” si se interconectan con el RI de la institución a través de protocolos de interoperabilidad.

Así bien, una institución que cuenta con un RI puede integrar uno o más repositorios federados que le brindan portales externos obteniendo como resultado la suma de la producción intelectual de su comunidad que fue publicada en la vía dorada del OA (figura 5).

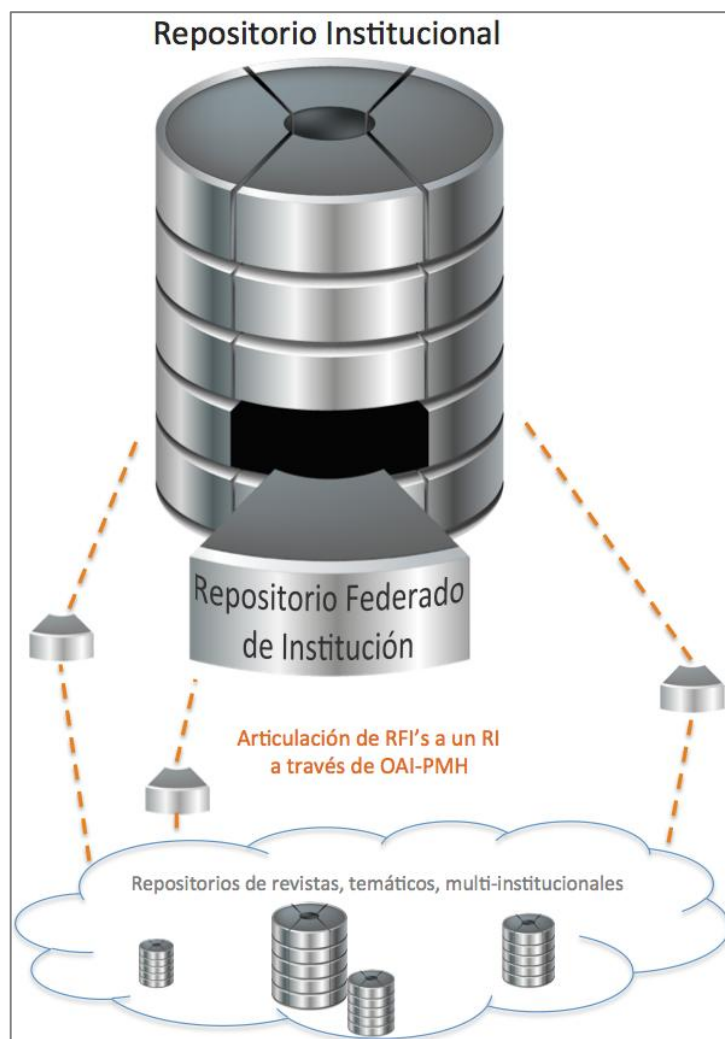


Figura 5. Repositorios Federados de Institución.

Entonces, los “Repositorios Federados de Institución” (RFI) se definen como repositorios externos interoperables que contienen la producción científica de los autores adscritos a la institución y que fue publicada en revistas de acceso abierto. Un

RFI es para una institución un micrositio para el acceso, búsqueda y consulta de los textos correspondientes a la producción científica de sus investigadores, el complemento a su RI para la producción científica publicada en revistas arbitradas de OA que puede ser integrado a través de protocolos de interoperabilidad como OAI-PMH.

Si bien el incremento en la visibilidad del conocimiento científico fue el objetivo inicial de los repositorios y portales de revistas arbitradas históricamente hasta fines de los 90; hoy en día la vía dorada y sus repositorios han alcanzado a cubrir otras funciones, como:

1. Organizar y clasificar el conocimiento.
2. Validar y/o certificar la calidad de lo publicado.
3. Generar indicadores de medición de los procesos de comunicación científica.
4. Convertirse en fuentes de información para investigación.
5. Fungir como espacios de intermediación para habilitar nuevas tecnologías como por ejemplo: *open data*, archivos abiertos, etc.
6. Gestionar formas de interoperabilidad y preservación.
7. Maximizar el aprovechamiento de la información por entidades, es decir, para autores, instituciones, países, etc. (Becerril-García, Aguado-López, Rogel-Salazar, Garduño-Oropeza, & Zúñiga-Roca, 2012).

Un RFI es generado a partir de un portal con acervo de múltiples publicaciones por lo que además aporta los siguientes valores agregados a una institución:

- Maximiza la visibilidad e impacto potencial de los artículos a través de herramientas de interoperabilidad.

- Genera indicadores cienciométricos de producción, colaboración o uso, lo que permite la comparación y caracterización de la producción científica en un contexto temático o regional.

Con esta propuesta los portales de revistas, repositorios multi/cross-institucionales o repositorios temáticos se convierten en potenciales fuentes generadoras de repositorios federados para las instituciones, siempre y cuando:

1. Permitan la agrupación de artículos basada en los datos de afiliación institucional de los autores.
2. Implementen un proceso de normalización de los datos de afiliación institucional.
3. Pongan a disposición dichas agrupaciones de artículos a través de OAI-PMH cosechables a nivel institución.

Es decir, portales como Scielo, Redalyc, ArXiv, RePEc, y muchos otros, son potenciales *pobladores* de los repositorios institucionales. *Pobladores* en el sentido de transferir a través de protocolos de interoperabilidad los metadatos de sus artículos a las instituciones que los produjeron, apalancando a su vez redes nacionales (como REMERI) o regionales (como LaReferencia).

Con ello, no solo se evita la duplicación de esfuerzos y almacénes, también se acelera la obtención de los beneficios de los RI's para las instituciones. No hay duda que la institución debería archivar sus contenidos; sin embargo, la realidad es que muchos de esos contenidos están ya archivados en repositorios de revistas arbitradas. ¿Por qué no *devolver* a las instituciones su producción y acelerar la población de los RI's en beneficio de la visibilidad y el acceso a los contenidos?. El horizonte de la comunicación científica es el *Open Access* (OA), de lo que se trata es de acelerar la transición. (Finch, 2012)

No debe olvidarse que la razón de ser de un repositorio institucional es devolver valor a la institución o comunidad que lo soporta, buscando la maximización de la visibilidad y el impacto del conocimiento (Swan, 2012). En este sentido, los RFI retornarían ese valor a las instituciones a través de la disponibilidad en línea de su producción científica de la vía dorada del acceso abierto.

#### 4. Los Repositorios Federados de Institución en Redalyc

Desde la nueva versión del portal Redalyc.org (liberada en Enero de 2013) es posible encontrar micrositios de más de 15,000 instituciones con la producción científica de sus autores (Redalyc, Red de Revistas Científicas de América Latina y el Caribe), de las cuales 9,908 son instituciones de América Latina y el Caribe.

En la siguiente figura se muestra el ejemplo del micrositio de la Universidad Nacional Autónoma de México (UNAM) en Redalyc.org, que contiene más de 10,000 trabajos de investigadores adscritos a la universidad, mismos que pueden ser recuperados para su consulta a través de motores de búsqueda. Este micrositio es a su vez un “Repositorio Federado” para la UNAM ya que implementa el protocolo OAI-PMH y puede ser integrado en el repositorio de dicha institución.

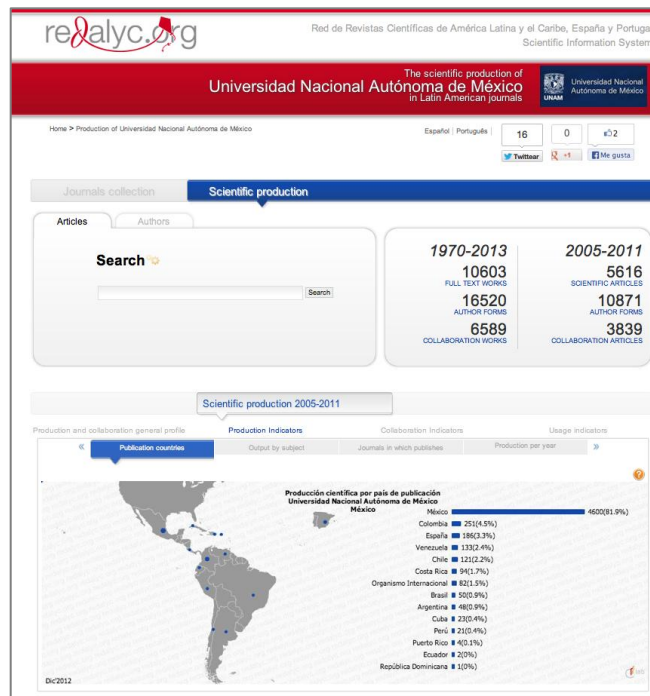


Figura 6. “Repositorio Federado” de la UNAM en Redalyc.org (**Redalyc, La producción científica de la UNAM en revistas indizadas en Redalyc.org**)

El acervo de redalyc que asciende actualmente a 284,159 artículos a texto completo es procesado a través de procedimientos de normalización y minería de datos para poder agrupar los trabajos según la institución generadora de cada uno, contemplando factores como diferentes formas de autor, formas de institución, datos incompletos, entre otros.

En la siguiente gráfica se observa la producción científica de las 10 instituciones más productivas de ALyC publicada en revistas indizadas en Redalyc, con ello se deja ver la potencialidad y relevancia de esta base de datos para *retornar* a la institución sus artículos publicados en revistas científicas.

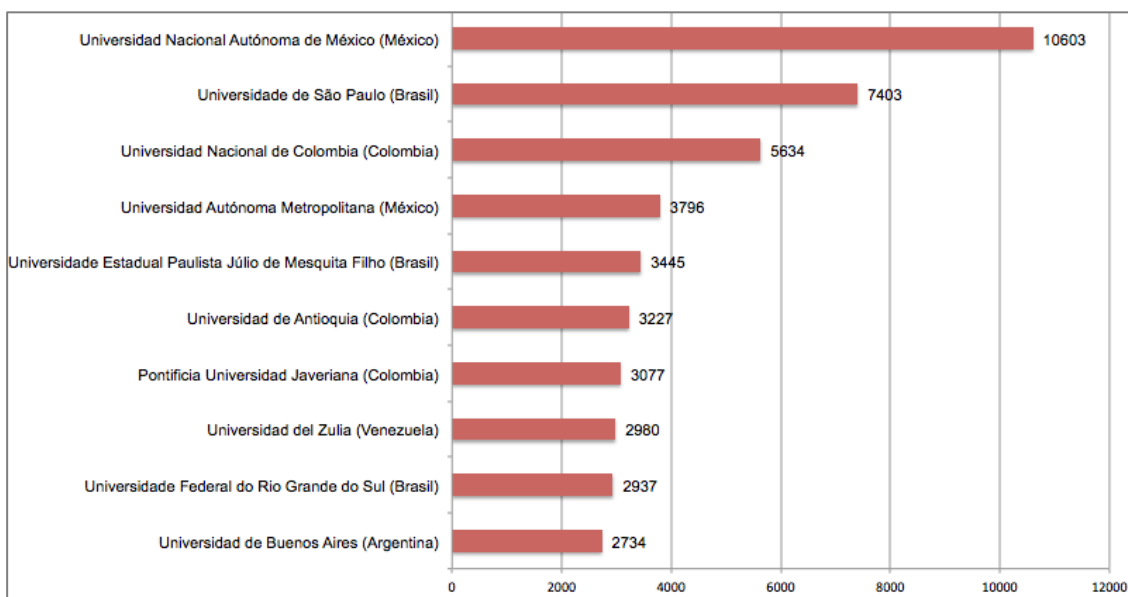


Figura 7. Producción por institución publicada en revistas de Redalyc. Instituciones más productivas de ALyC

## 5. Conclusiones

La generación de Repositorios Federados de Institución por parte de los portales o repositorios de revistas arbitradas de acceso abierto, repositorios temáticos cross/multi-institucionales permitirían poblar los RI con los metadatos de los artículos publicados en la vía dorada del acceso abierto.

Con ello una institución integraría los contenidos que hasta ahora, por circunstancias de transición entre el modelo tradicional de publicación y los modelos OA que permiten las nuevas tecnologías, no han permitido a los RI's llegar a su máximo potencial.

La articulación a los repositorios institucionales de servicios interoperables que provean la producción científica publicada en revistas arbitradas de acceso abierto enriquece y recupera los contenidos que hasta ahora han permanecido inalcanzables para el auto-archivo en las instituciones.



La propuesta concreta se materializa en los RFI's generados por Redalyc.org, y adquieren relevancia dadas las condiciones presentes en América Latina y el Caribe. Su uso e integración potencializa los beneficios de la interoperabilidad, el incremento en la visibilidad, uso e impacto; pero sobretodo, es un medio de devolver a las instituciones la posibilidad de reunir el conocimiento generado por su comunidad.

## 5. Bibliografía

Ware, M. (2012). *The STM report: an overview of scientific and scholarly journals publishing* (3a ed.). Oxford, STM.

Abadal, E. (Mayo de 2013). Hacia dónde van las revistas científicas. *blok de bid*.

Becerril-García, A., Aguado-López, E., Rogel-Salazar, R., Garduño-Oropeza, G., & Zúñiga-Roca, M. (2012). De un modelo centrado en la revista a un modelo centrado en entidades: la publicación y producción científica en la nueva plataforma Redalyc.org. (I. U. Oviedo, Ed.) *Aula Abierta*, 40 (2), 53-64.

Canessa, E., & Zennaro, M. (2008). *Science Dissemination using Open Access* (1a ed.). ICTP - The Abdus Salam International Centre for Theoretical Physics.

Finch, D. (2012). *Accessibility, sustainability, excellence: how to expand access to research publications*. Report of the Working Group on Expanding Access to Published Research Findings, Finch Group.

Hubbard, B. (2012). *Green, Blue, Yellow, White & Gold. A brief guide to the open access rainbow*. (R. S. Project, Productor) Recuperado el 5 de Mayo de 2013, de Sherpa / RoMEO:  
<http://www.google.com.mx/url?sa=t&rct=j&q=&esrc=s&source=web&cd=5&cad=rja&ved=0CG8QFjAE&url=http%3A%2F%2Fwww.sherpa.ac.uk%2Fdocuments%2Fsherpaplusdocs%2FNottingham-colour->

guide.pdf&ei=Hji6UaOjGaW\_0gGHhoHQBQ&usg=AFQjCNFr1KMdAvxCh9aopreOCz-WRjf27g&sig2=7J6fpJIXafaXWjiZp2TPmg&bvm=bv.47883778,d.dmQ

Redalyc. (s.f.). *La producción científica de la UNAM en revistas indizadas en Redalyc.org*. Recuperado el 6 de Junio de 2013, de Redalyc.org:  
<http://www.redalyc.org/institucion.oa?id=15992&tipo=produccion>

Redalyc. (s.f.). *Red de Revistas Científicas de América Latina y el Caribe*. Recuperado el 12 de Mayo de 2013, de [www.redalyc.org](http://www.redalyc.org).

REMERI. (2012). *Red Mexicana de Repositorios Institucionales*. Recuperado el 15 de Junio de 2013, de [www.remeri.org.mx](http://www.remeri.org.mx)

ROAR. (s.f.). *Registry of Open Access Repositories*. Recuperado el 2 de Junio de 2013, de [roar.eprints.org](http://roar.eprints.org)

Rodríguez, T. M. (Enero de 2013). *Diagnóstico para crear la Red Mexicana de Repositorios Institucionales*. Recuperado el 6 de Junio de 2013, de [www.remeri.org.mx](http://www.remeri.org.mx): <http://www.remeri.org.mx/remeri/pproyecto.html#diagnostico>